energies

MDPI

Article

Feasibility of Harris Hawks Optimization in Combination with Fuzzy Inference System Predicting Heating Load Energy Inside Buildings

Hossein Moayedi [1,2,*] and Bao Le Van [1,2]

[1] Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam
[2] School of Engineering and & Technology, Duy Tan University, Da Nang 550000, Vietnam
* Correspondence: hosseinmoayedi@duytan.edu.vn

Abstract: Heating and cooling systems account for a considerable portion of the energy consumed for domestic reasons in Europe. Burning fossil fuels is the main way to produce this energy, which has a detrimental effect on the environment. It is essential to consider a building's characteristics when determining how much heating and cooling is necessary. As a result, a study of the related buildings' characteristics, such as the type of cooling and heating systems required for maintaining appropriate indoor air conditions, can help in the design and construction of energy-efficient buildings. Numerous studies have used machine learning to predict cooling and heating systems based on variables that include relative compactness, orientation, overall height, roof area, wall area, surface area, glazing area, and glazing area distribution. Fuzzy logic, however, is not used in any of these methods. In this article, we study a fuzzy logic approach, i.e., HHO−ANFIS (combination of Harris hawks optimization and adaptive neuro-fuzzy interface system), to predict the heating load in residential buildings and investigate the feasibility of this technique in predicting the heating load. Fuzzy techniques obtain perfect results. The analysis results show that the HHO−ANFIS with a population size of 400, the highest value of R2 (0.98709 and 0.98794), and the lowest value of RMSE (0.08769 and 0.08281) in the training and testing dataset, respectively, can predict the heating load with high accuracy. According to the high value of R2 (98%) and low value of RMSE, HHO−ANFIS can be used in predicting the heating load of residential buildings.

Keywords: ANFIS; heating-load; metaheuristic; residential buildings

check for updates

Citation: Moayedi, H.; Le Van, B. Feasibility of Harris Hawks Optimization in Combination with Fuzzy Inference System Predicting Heating Load Energy Inside Buildings. *Energies* **2022**, *15*, 9187. https://doi.org/10.3390/en15239187

Academic Editor: Chi-Ming Lai

Received: 30 September 2022
Accepted: 2 December 2022
Published: 4 December 2022

1. Introduction

Globally, cities and the population of people are growing rapidly, and meeting citizens' needs requires a lot of energy. According to recent studies, the number of people living in cities is predicted to grow to five billion at the end of 2030 [1]. Dwellings use a significant amount of energy, and other types of buildings comprise only a small portion of the total energy consumption [2,3]. There are limited resources available for supplying energy to citizens. Due to the importance of residential buildings to the overall welfare of society, residential consumption needs to be carefully monitored and controlled [4,5]. Meanwhile, cooling and heating systems are two essential energy resources for citizens, so their usage should be managed accordingly.

It is critical to have complete knowledge about buildings' performance and environmental factors to manage and optimize their energy consumption. In a building, electricity, heating supply, and gas are the primary energy sources, but domestic hot water, heating, ventilation and air conditioning (HVAC), elevators, and other applications are the essential final uses. Building energy performance is influenced by all these mentioned resources, but two are the most critical factors: optimum indoor air condition supply operation and HVAC systems [6,7]. HVAC plays a vital role in residential buildings by regulating the internal cooling and heating loads [8]. In spite of the necessity of this system for buildings, it consumes about 40% of the overall energy, particularly in office buildings [9,10]. In order to minimize the heating and cooling costs of buildings, it is essential to forecast the thermal loads because deviations from the optimal scheduled values will greatly increment the whole costs [11].

A residential building's consumption patterns can be predicted with the help of energy forecasting to maintain an optimal HVAC and energy management system [12,13]. The development of technology has enabled many micro-intelligent building management systems (BMSs) and devices to be installed on green building sites, which record and monitor building load patterns that influence energy forecasting. Such data could be used to predict and control building consumption patterns on an hourly basis. A number of studies have explained the importance of forecasting energy. In [14], for example, a survey of thermal energy consumption in buildings was conducted. This study investigated fuel mix, social–economic conditions, and climate change as factors that affect thermal energy comfort. As a result of forecasting the loads, buildings can flexibly schedule energy consumption for the following day, not only in order to participate in demand-response programs [15–17], but also for energy trading programs [18,19].

Building load forecasting has been evaluated in a number of studies so far. Authors in Ref [20], used an integrated design method to estimate energy savings over the lifetime of a building, minimize expenses, and carbon emissions reductions. Applying a genetic algorithm integrated with a dynamic simulating tool, [21] presents a multi-objective optimization technique for renovating existing buildings and HVAC systems. There is an investigation of the effect of data dimensionality and size on artificial neural networks (ANNs) in [22–26], where the predictive power of ANNs is evaluated for office buildings. A study using ANN has been conducted in [27] to forecast electricity load for HVAC systems. The BR-based ANN outperforms the others among the three algorithms applied in this paper, including scaled conjugate gradient backpropagation, Bayesian regularization (BR), and Levenberg–Marquardt. In another research, statistical analysis was used to forecast cooling and heating in an office [28]. Moreover, In Ref. [29], four hybrid methods based on metaheuristics and ANN have been presented for forecasting buildings' energy efficiency based on particle swarm optimization (PSO), artificial bee colony (ABC), genetic algorithms (GA), and imperialist competitive algorithms (ICA). According to [30], cooling load forecasting was achieved using probabilistic entropy-based neural networks (PENNs). The energy consumption of a building was reduced by 36.5% through the use of feed forward neural networks (FFNNs) in [31]. A decision tree approach has been proposed in [32] for predicting energy requirements and evaluating energy performance measures for residential buildings. An analysis of forecasting strategies for cooling and heating loads was conducted in [33], comparing machine learning (ML) approaches such as deep neural network (DNN) [34], Gaussian process regressions (GPR), gradient boosted machines (GBM), and minimax probability machine regressions (MPMR). Furthermore, according to [35], ANN, general linear regression (GLR), classification and regression tree (CART), and chi-squared automatic interaction detectors (CHAID) were applied to predict the cooling and heating systems of buildings. This paper considers the building's technical characteristics as an input for the networks. Adaptive linear time-series methods and models were used to forecast cooling and heating energy consumption for sixteen residential buildings in Ref [36]. Similarly, [37] presented a BMS that can forecast cooling loads with the help of data mining techniques [38]. Optimization of HVAC heat storage in public buildings has also been investigated with general regression neural networks (GRNNs) in [39].

The application of fuzzy logic methods for energy analysis in buildings has mainly taken place in control research [40,41] and multi-criteria decision-making [42,43]. However, as far as the authors are aware, fuzzy logic has only been used sparingly in energy performance estimation.

An adaptive neuro-fuzzy inference system (ANFIS) model is proposed in this research for predicting residential energy performance. An ANFIS modeling approach is a very good option in the energy management field of buildings due to its potential for working with uncertainties, explaining effective complex relationships, and identifying uncertain causal relationships between variables [44]. Moreover, its ability to accurately predict energy consumption is extremely valuable [45,46]. As a consequence, we explore non-linear fuzzy approaches for estimating residential buildings' energy performance, including fuzzy inductive reasoning (FIR) and ANFIS. In Ref [47], a first attempt has been made in this direction.

Xifara and Tsanas [48] have created simulated buildings of the residential type that are accessible in the UCI machine-learning repository [49]. Using simulated building data to deal with energy prediction problems is an important and justified method, since it reduces the amount of time required to achieve the data, allows researchers to design various building configurations easily, and allows them to work with lower detail levels [50]. Based on the authors' assertions, the produced data is likely to represent actual real data, allowing for comparisons of energy between buildings.

In their paper, residential buildings' cooling and heating loads were predicted using classical linear regression and random forest approaches [48]. Similar problem and datasets were addressed by other studies using additional ML approaches, such as support vector machines, neural networks, and evolutionary computation [34,35,51–55].

Having studied the heating load in residential buildings using HHO integrated with ANFIS, this paper's first goal is to compare its performance. With the mentioned model, the heating load can be forecasted. The contribution of the current study is to exert a hybrid method of artificial intelligence, for evaluating the heating load. Several buildings (768) are investigated for this study. This information is then trained utilizing the HHO−ANFIS. The outcomes of HHO−MLP are reported utilizing three criteria of performance and recognize the feasibility and accuracy level of the method in forecasting residential buildings' heating load. It is a novelty to investigate the application of fuzzy approaches for the prediction of heating and cooling loads based on building design characteristics because no other study has used this type of system.

The remainder of the article is classified as follows. The datasets used for this paper are extensively discussed in Section 2. Section 3 introduces HHO and ANFIS methodologies and describes how the HHO−ANFIS model forecasts the energy performance in residential buildings. In Section 4, a discussion of the results achieved by the HHO−ANFIS approach is presented, and its effectiveness is demonstrated compared to other strategies. Finally, the conclusion is presented in Section 5.

2. Established Database

Our study utilized data from the UCI machine learning repository [49] with the following:

A total of 768 simulated buildings were generated using Ecotet by the authors in Ref [48]. The Ecotet is a special software tool for sustainable building design. It comprehensively analyzes a building's energy, water use, and thermal performance, among other capabilities [56].

Each building in the simulation occupies a total of 771.75 m3, but each has a different surface area and dimensions. Building elements are designed to achieve the minimum heat loss (U-value) from walls, floors, or roofs using the latest and most frequently used materials in the building industry. In this simulation, buildings are assumed to be residential buildings in Athens, Greece.

Three categories of glazing areas were applied as a percentage of the floor area: 10%, 25%, and 40%. In addition, five different scenarios were used to simulate the distribution of each glazing area: (1) uniform scenario: 25% glazing area for all parts, (2) north scenario: 55% for the north and 15% for any other parts, (3) east scenario: 55% for the east part and 15% for any other parts, (4) south scenario: 55% for the south part and 15% for any other parts, and (5) west scenario: 55% for the west part and 15% for any other parts. Furthermore, samples without glazing areas were obtained. A final rotation was performed to ensure that all shapes were facing the four fundamental points. Figure 1 shows the building's types.

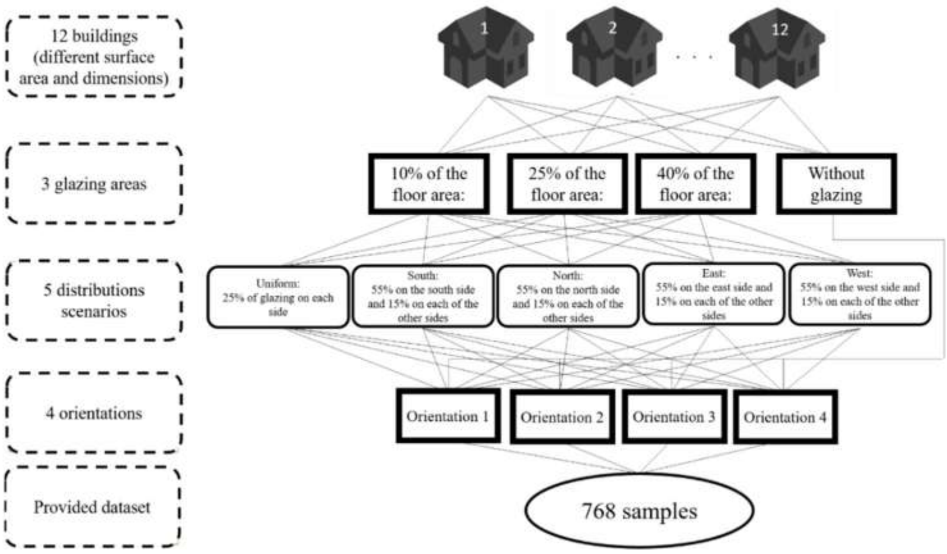


Figure 1. Graphical view of data preparation.

Table 1 summarizes eight building parameters that characterize all 768 simulated buildings. Parameters corresponding to input variables are listed in this table.

Table 1. Input variables used for the models.

RC	Relative Compactness	OH	Overall Height
SA	Surface Area	**O**	Orientation
WA	Wall Area	**GA**	Glazing Area
RA	Roof Area	**GAD**	Glazing Area Distribution

Furthermore, the output variable of heating load (HL) was recorded. The data were also analyzed statistically, and linear modeling approaches were not suitable because the density and scatter plots were not according to Gaussian distributions. It has already been mentioned that this data is likely to represent real data, allowing energy comparisons among buildings [48]. As an additional benefit, simulated building data can be very beneficial and justified when it comes to energy prediction issues since it reduces the time it takes to obtain data, allows us to design different building configurations easily, and allows us to work with lower degrees of detail during the prediction process [57].

3. Methodology

The methodology of paper searching is to predict the heating load in residential buildings via the fuzzy approach of HHO−ANFIS. The following subsections introduced both the HHO and ANFIS approach widely, and in the next section the analysis used criteria in forecasting the heating load, and the results are discussed.

3.1. Adaptive Neuro-Fuzzy Interface System (ANFIS)

In terms of function approximation, ANFIS proposed by Jang is a known strategy among hybrid neuro-fuzzy systems with several applications in different fields [58,59]. The ANFIS is a neuro-fuzzy system of the Sugeno type [60]. Using neural networks as a learning method, neuro-fuzzy systems find their parameters through a learning process [61–63]. In line with the HHO methodology, the learning technique is data-driven, not knowledge-based [64,65].

As expressed in Equation (1), one of the main characteristics of the Sugeno inference system is that the outputs of rules in the fuzzy system are functions rather than fuzzy sets.

$$\begin{array}{l} \mathrm{R1} = \text{If } a \text{ is } A_1 \text{ and } b \text{ is } B_1 \text{ then } z = p_1 \times a + q_1 \times b + r_1 \\ \mathrm{R1} = \text{If } a \text{ is } A_2 \text{ and } b \text{ is } B_2 \text{ then } z = p_2 \times a + q_2 \times b + r_2 \end{array} \quad (1)$$

In this equation, b and a are linguistic variables; B_2 , B_1, A_2, and A_1 are linguistic values selected by fuzzy sets on the scope of B and A; p_1, q_1, r_1, p_2, q_2, and r_2 are output function variables.

In Figure 2, the two rules expressed in Equation (1) are illustrated graphically to explain the inference process based on the Sugeno model. To obtain the degree of membership of each input to each fuzzy set, the Sugeno inference first combines a given input tuple (for this figure: a = 3 and b = 2) with its antecedents (left panel of Figure 2). To compute the final output or z (right panel of Figure 2), the main function is applied to get the weight for each rule or w_i. Two different sets of parameters are involved in the Sugeno inference. The first set relates to the input variables and their membership function parameters. Additionally, there is a second set containing the parameters related to the output function for each rule, which are p_i, q_i, and r_i.

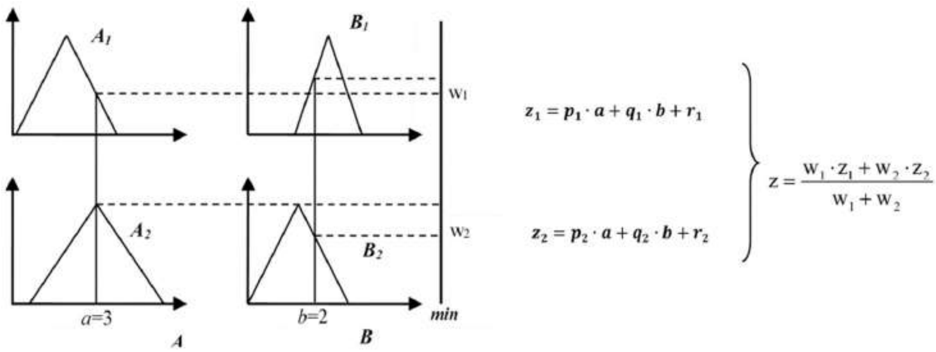


Figure 2. Example of a Sugeno model (estimation of two fuzzy rules with two input variables, b = 2 and a = 3).

The parameters in these two sets are automatically adjusted by ANFIS using least square and backpropagation (gradient descendent) estimation algorithms [66]. For learning the variables of the antecedents or membership functions, backpropagation can be applied, while least squares estimation determines the coefficients for the linear combinations to obtain the rules' results. Figure 3 shows the ANFIS structure, which consists of five layers. In ANFIS, n represents the number of input nodes. For the example in Figure 3, the value of n is 2 and is related to a and b.

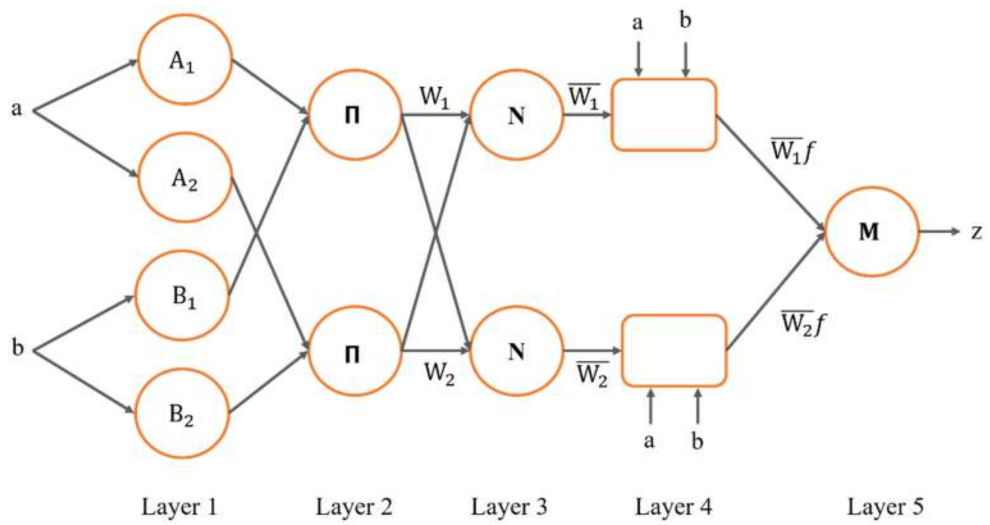


Figure 3. Adaptive neuro−fuzzy inference system (ANFIS) architecture.

Each node in layer 1, the fuzzification layer, represents a membership value to a linguistic term, such as a triangular, Gaussian, or other function. Backpropagation is used during the learning process to adjust the parameters for the chosen function.

When the parameters change, so do the linguistic terms of A_i, and B_i in the membership function. The fuzzy intersection is performed at layer 2 by using the T-norm operator. In addition, in layer 3, each rule is normalized, and its strength is calculated using $\overline{\omega}_i = \frac{\omega_i}{\sum_i \omega_i}$. Moreover, in layer 4, an algorithm based on least squares approximation to update the parameters of the results of the rules. As a final step, layer 5 sums up the outputs of all layer 4 nodes. In the MATLAB program, ANFIS is implemented as a function in the Fuzzy toolbox. In Nauck et al. [58], the ANFIS methodology is explained in more detail.

3.2. Harris Hawks Optimization (HHO)

In their study [67], Heidari et al. developed a new optimization method called HHO. This nature-inspired optimization algorithm is based on the behavior model of Harris Hawk birds. Cooperation between hawks to hunt prey is a vital part of the algorithm. In order to take the prey by surprise, a flock of Harris hawks attacks from different directions based on the algorithm. The Harris hawk's chase model correlates with the prey's escape pattern. In the attacking process, birds cooperate. Meanwhile, the Harris hawk leading the hunt attacks the prey, follows, and unexpectedly leaves out of sight, but the next Harris hawk stays on the chase. By using this strategy, the prey is eventually exhausted and captured. Compared to other algorithms, HHO is better suited for constraint-based problems. In addition, as a global optimizer, HHO can maintain a balance between the exploration and exploitation phases [68]. In this optimization algorithm, three phases are involved. Exploration is the first phase, and it can be described using the following equation:

$$x(t+1) = \begin{cases} x_{rand}(t) - r_1|x_{rand}(t) - 2r_2x(t)| & q \geq 0.5 \\ x_{prey}(t) - x_a(t) - r_3(LB + r_4(UB - LB)) & q < 0.5 \end{cases} \tag{2}$$

where $x(t)$ represents the current position of Hawk, $x(t+1)$ represents Hawk's position in the succeeding iteration t, $x_{prey}(t)$ represents the position of prey, Each of r_1, r_2, r_3, r_4 and q is a randomly chosen number between 0 and 1. $x_{rand}(t)$ represents a randomly chosen hawk among the population. Moreover, LB and UB are the lower and upper limits, respectively. $x_a(t)$ represents the average position of the Harris Hawk, which can be obtained by:

$$x_a(t) = \frac{1}{N}\sum_{i=1}^{N} x_i(t) \tag{3}$$

In this equation, $x_i(t)$ is the position of each Harris Hawk in iteration number t, and N represents the overall number of Harris Hawks. Exploitation is the second phase. The energy of hawks is reduced through the hunt-and-chase process. Preys' energy can be formulated by:

$$E = 2E_0\left(1 - \frac{1}{T}\right) \tag{4}$$

where E represents escaping energy, E_0 is the amount of energy in the initial stage, and T is the maximum allowable number of iterations. For this phase, when $|E_0| \geq 1$ exploration occurs and when $|E_0| < 1$ exploitation happens.

Exploitation is the third phase, which involves improving locally found solutions using previously collected data. During this phase, the hawks surprise their prey, identified in the second phase, by attacking it. Four models have been presented for the phase of the attack, formulated on the prey's escape and the hawks' chasing behavior.

3.2.1. Soft Besiege

Condition for soft besiege strategy is justified when $r \geq 0$ and $|E| \geq 0$, which can be modeled by:

$$x(t+1) = \Delta x(t) - E\left|Jx_{prey}(t) - 2x(t)\right| \tag{5}$$

$$\Delta x(t) = x_{prey}(t) - x(t) \tag{6}$$

In this equation, Δx represents the difference between the prey and Hawk's current position in the t iteration. J is the prey jump power parameter in the escaping process, which is calculated by $J = 2(1 - r_5)$ and r_5 considered as a randomly selected number between 0 and 1.

3.2.2. Hard Besiege

The condition for hard besiege strategy is justified when $r \geq 0$ and $|E| < 0$. The prey does not have enough energy to escape because it is tired. This phase can be modeled by:

$$x(t+1) = x_{prey}(t) - E_n|\Delta x(t)| \tag{7}$$

3.2.3. Soft Besiege with Progressive Rapid Dive

The condition for soft besiege with progressive rapid dive strategy is justified when $r < 0$ and $|E| \geq 0$. The prey is able to escape successfully during this phase. For implementing a soft besiege, Hawk considers the following next move:

$$x = x_{prey}(t) - E\left|Jx_{prey}(t) - x(t)\right| \tag{8}$$

$$Z = Y + S + LF(D) \tag{9}$$

In this equation, D represents the dimension and S are considered as a random vector of $1 \times D$ size, and LF represents the levy flight function [67,69]. Consequently:

$$x(t+1) = \begin{cases} Y & f(Y) < f(y(t)) \\ Z & f(Z) < f(y(t)) \end{cases} \tag{10}$$

Hard besiege with progressive rapid dive:

The condition for a hard besiege with progressive rapid dive strategy is justified when $r < 0$ and $|E| < 0$. In this case, it is impossible for the prey to escape appropriately because it lacks sufficient energy. Using the following equation, this strategy can be formulated:

$$x(t+1) = \begin{cases} x_{prey}(t) - E\left|Jx_{prey}(t) - x_m(t)\right| & f(Y) < f(y(t)) \\ Z = Y + S + LF(D) & f(Z) < f(y(t)) \end{cases} \tag{11}$$

4. Results and Discussion

A two-stage computational model is applied in this paper to predict heating load (HL) using influential variables. The training set comprises almost 80% of the samples, and the testing set comprises approximately 20%. This ensures the generalization capabilities of the model and overcomes the overfitting problem.

4.1. Accuracy Indicators

A training subset is used to obtain model parameters, and a testing subset is used to validate the model. Statistical confidence is achieved by repeating the training and testing phases 1000 times on the whole dataset, rearranging them randomly in each run before segmenting them into training and testing.

Each of the models is evaluated according to three criteria. Two of these criteria are error indexes, including the root mean square error (RMSE) and the mean absolute error (MAE), which are formulated in Equations (12) and (13), respectively. The indexes selected here are those utilized most frequently in previous research works [35,48,52], so our proposed fuzzy technique can be compared with their methodologies.

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|^2} \quad (12)$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \quad (13)$$

In these equations, $\hat{y}(t)$ refers to the predicted output, $y(t)$ is the actual output, and N represents the overall number of samples. This study compares various strategies on the same dataset using RMSE and MAE, which are commonly used as accuracy indexes. All individual differences in the MAE are weighted equally in the average since it is a linear index. In the RMSE, errors are squared before they are averaged, which provides a quite high weight for large errors. In other words, it is most beneficial to use the RMSE when large errors are especially undesired. It is possible to reduce the variation in forecast errors by combining the MAE and the RMSE.

To have a comprehensive measure of performance, the third evaluation criterion was implemented. For this purpose,

$$\mathrm{R}^2 = 1 - \sum_{i=1}^{N}(y_i - \hat{y}_i)^2 / \sum_{i=1}^{N}(y_i - \overline{y}_i)^2 \quad (14)$$

The term $\overline{y}_i$ stand for the number of occurrences and the average of the actual HL values.

4.2. Incorporated FIS with Optimizers

In this section, the RMSE and MAE values, calculated for HL output variables, are summarized. A diagram showing the accuracy obtained based on ideal values for each index can be found in this diagram. It should be noted that RMSE and R2 should ideally be 0 and 1. Based on a training population of 400 individuals, the values for these indexes related to the HHO−ANFIS model are 0.08769 and 0.98709. As a result, it can be concluded that the proposed model achieved 91.23% (1 − 0.08769 = 0.91231) and 98.70% (0.98709/1 = 0.98709) accuracy for RMSE and R2 indexes. This conclusion also applies to the validity and precision of the other indexes. It should be noted that to determine performance indexes in decimal form, the percentage values need to be converted. Figure 4 presents the MSE value versus iterations (1000 iterations) for ten population sizes (50, 100, 150, 200, 250, 300, 350, 400, 450, and 500) for HL in the present study. The best result is obtained from the lowest value of RMSE. As indicated in Figure 4 and Table 2, the lowest MSE value for HHO−ANFIS resulted from a population size of 400 (MSE = 0.08769 and 0.08281), and the highest MSE value from a population size of 500 (MSE = 0.15105 and 0.15016). These results indicate that the HHO−ANFIS can predict the heating load with high accuracy.

Zhang et al.'s [42] work introduced the theory of rank analysis. In each index, a maximum rank (which refers to the models' number under study) was assigned to the model that has the best value, whereas a rank of 1 was assigned to the model with the worst value, different for training and testing results. Their ranks were then added up to calculate the total score. As a final step, the testing and training phases are added together to calculate each model's final score.

Using the results demonstrated in Table 2, it is evident that the HHO−ANFIS approach made the most accurate prediction for predicting the heating load, applying a population of 400 and with the highest R2 values of 0.98709 and 0.98794 and the lowest RMSE values of 0.08768 and 0.08281, respectively, in the training and testing datasets. On the contrary, having the lowest R2 (0.96119 and 0.95978) and highest RMSE (0.15105 and 0.15016), the population of 500 is the least accurate model in heating load prediction. There is a strong correlation between the kind of input data used and the results of fuzzy methods. A comparison of the results obtained in this paper with those from other studies is necessary to analyze the efficacy of the suggested method. It is recommended to make comparisons using similar datasets with caution.

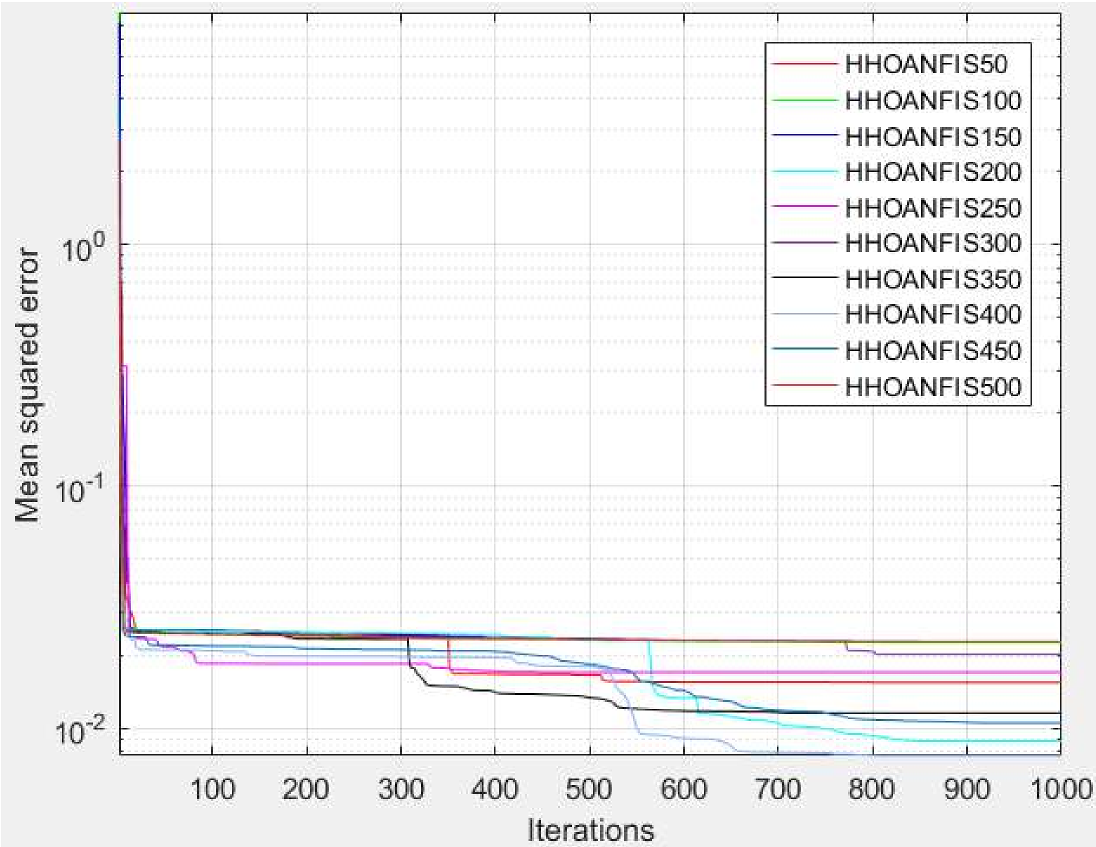


Figure 4. Variation of mean squared error versus iterations for the HHO−ANFIS.

Table 2. The network results for the HHO−ANFIS having different swarm size.

Swam Size	Training Dataset		Testing Dataset		Scoring				Total Score	Rank
	RMSE	R^2	RMSE	R^2	Training		Testing			
50	0.12403	0.974	0.126	0.97185	6	6	5	5	22	5
100	0.15039	0.96154	0.1469	0.96154	3	3	2	2	10	8
150	0.15102	0.96121	0.14546	0.96231	2	2	3	3	10	8
200	0.09207	0.98576	0.10153	0.98182	9	9	8	8	34	2
250	0.12921	0.97175	0.1248	0.97239	5	5	6	6	22	5
300	0.1409	0.96632	0.12994	0.97003	4	4	4	4	16	7
350	0.10745	0.98055	0.10199	0.98165	7	7	7	7	28	4
400	**0.08769**	**0.98709**	**0.08281**	**0.98794**	**10**	**10**	**10**	**10**	**40**	**1**
450	0.10233	0.98238	0.10016	0.98231	8	8	9	9	34	2
500	0.15105	0.96119	0.15016	0.95978	1	1	1	1	4	10

It is clear from Figure 5 that the proposed HHO−ANFIS network has passed its training phase properly since target data and network output correlate well. When training is carried out correctly, the network learns to identify patterns inherent in the data and then predicts unknown data by applying the learned patterns. Thus, the HHO−ANFIS network determines how much heating load is necessary for each building based on its particular characterizations. As a result of the training process, each network can predict the amount of heat load based on the input data from the test phase. The initial test data is used to validate each network after training. The network itself performs this test as part of the training phase.

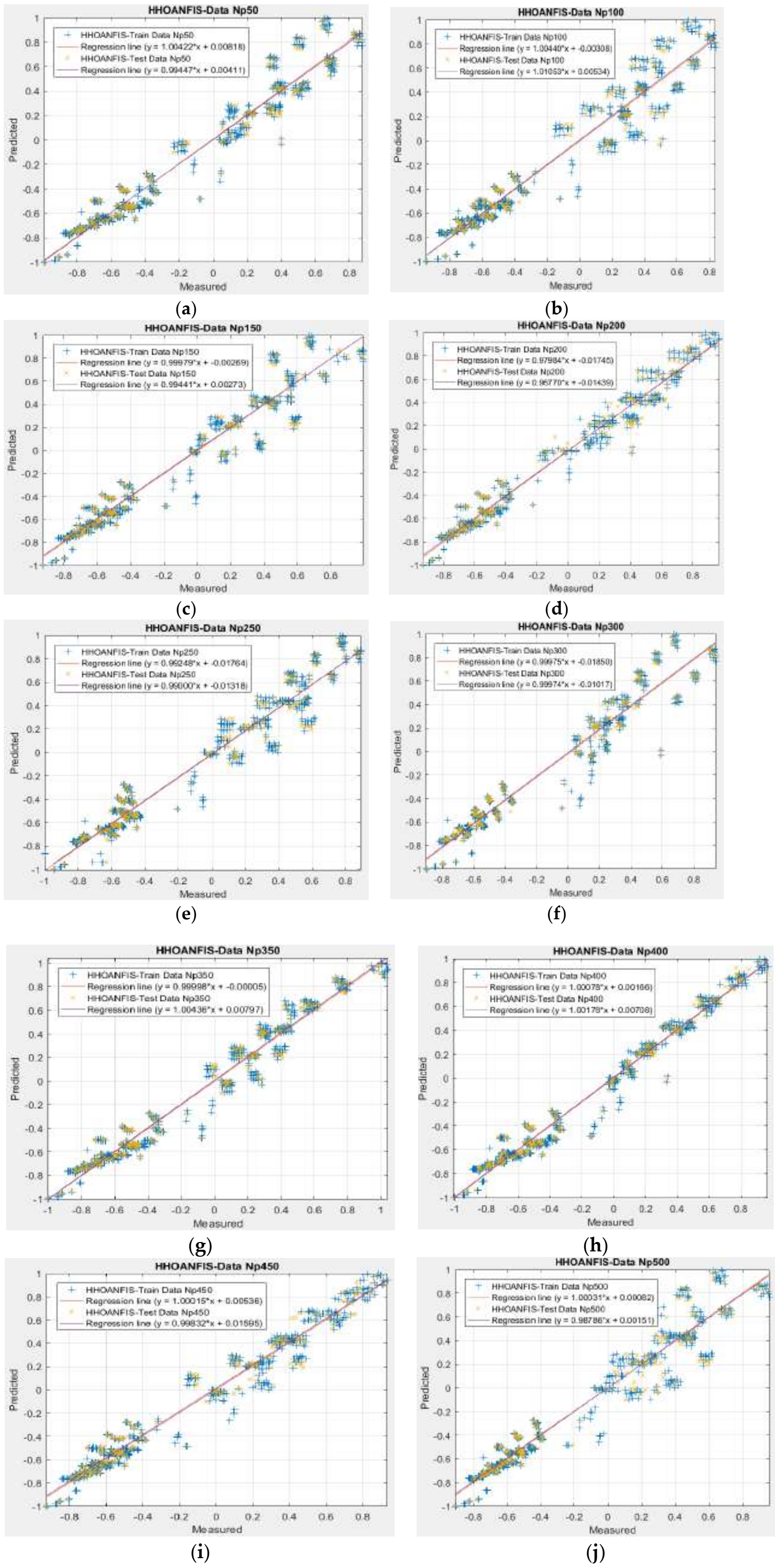
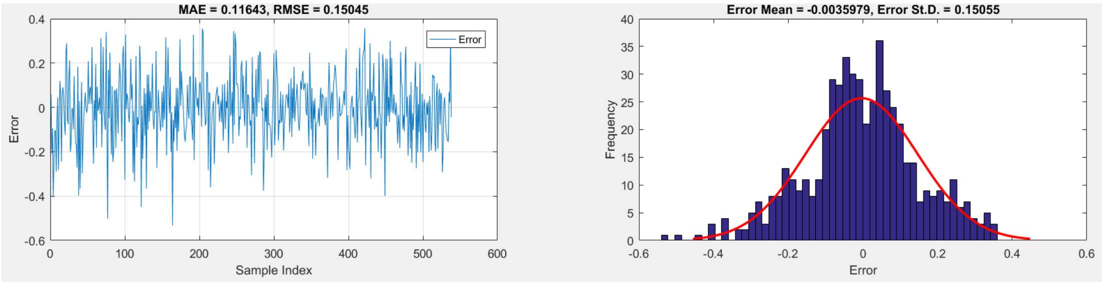


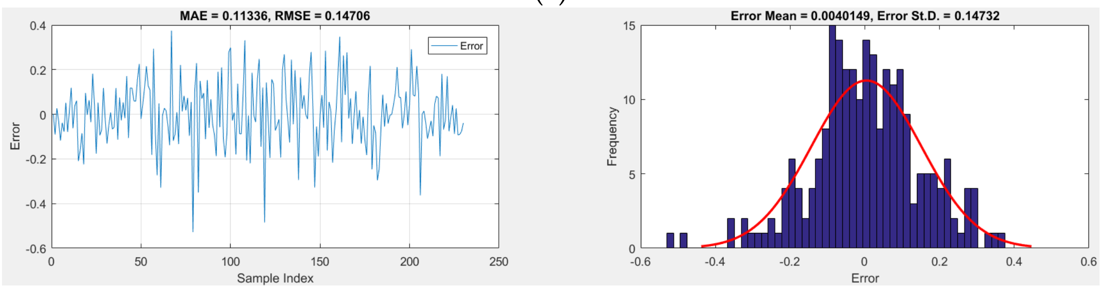
Figure 5. The accuracy of training and testing dataset performance of HHO−ANFIS in the best fit optimization structure. (**a**) HHO−ANFIS Np = 50. (**b**) HHO−ANFIS Np = 100. (**c**) HHOANFIS Np = 150. (**d**) HHO−ANFIS Np = 200. (**e**) HHO−ANFIS Np = 250. (**f**) HHO−ANFIS Np = 300. (**g**) HHO−ANFIS Np = 350. (**h**) HHO−ANFIS Np = 400. (**i**) HHO−ANFIS Np = 450. (**j**) HHO−ANFIS Np = 500.

4.3. Error Analysis

For the HHO−ANIFS network, the prediction errors for both the validation and test phases are depicted in Figures 6–10 in the histogram as one of the essential factors for assessing the results. This error histogram model indicates the maximum and minimum prediction errors. The trained network could have an amount of error equal to the values in the below figures when forecasting test data related to the heating load.

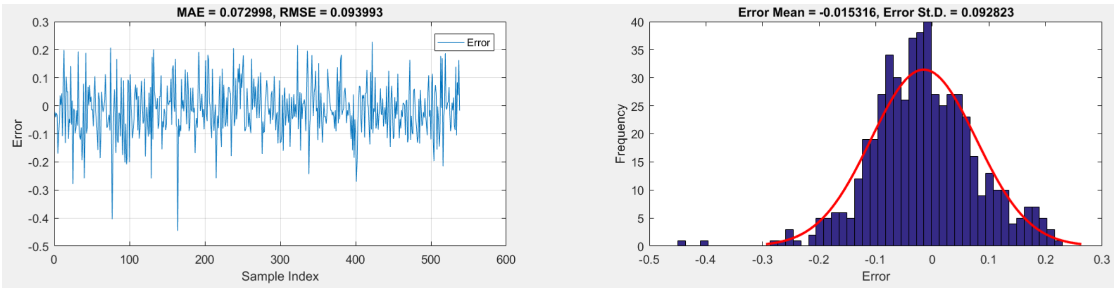


(**a**)

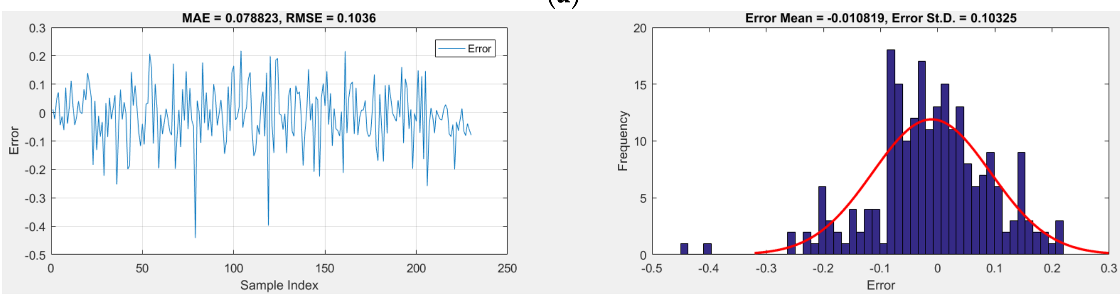


(**b**)

Figure 6. Frequency and minimum value of errors in HHO−ANFIS−100 best fit structure. (**a**) Training dataset. (**b**) Testing dataset.

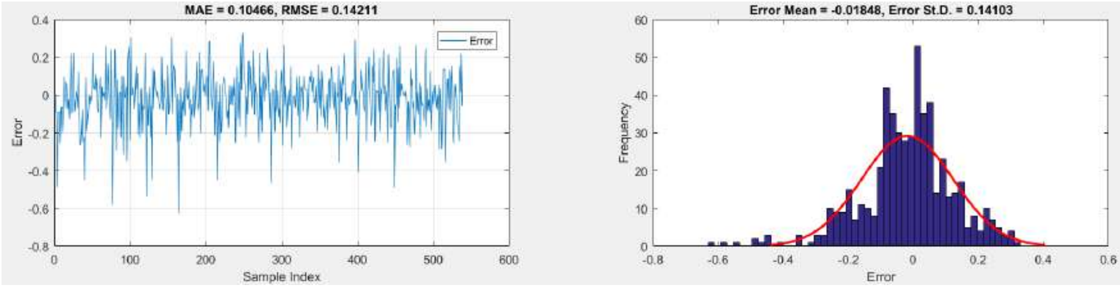


(**a**)

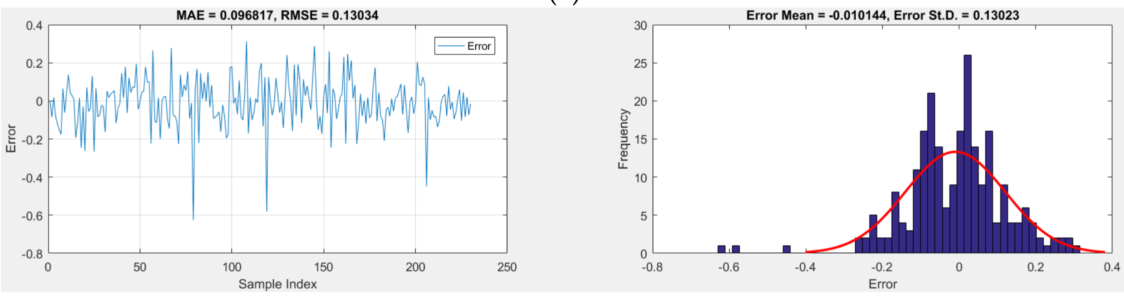


(**b**)

Figure 7. Frequency and minimum value of errors in HHO−ANFIS−200 best fit structure. (**a**) Training dataset. (**b**) Testing dataset.

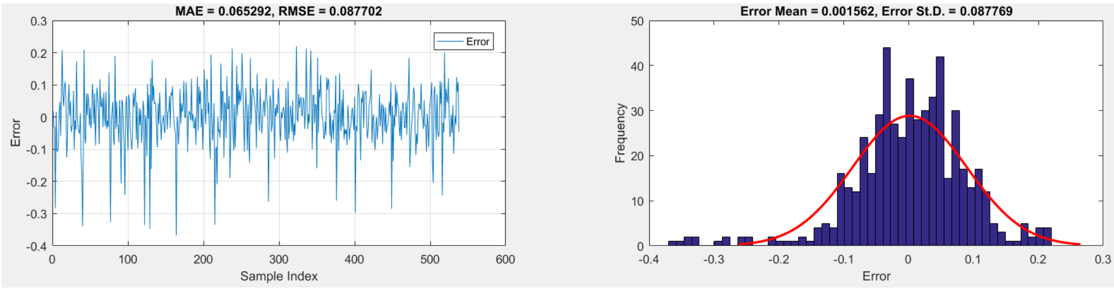


(**a**)

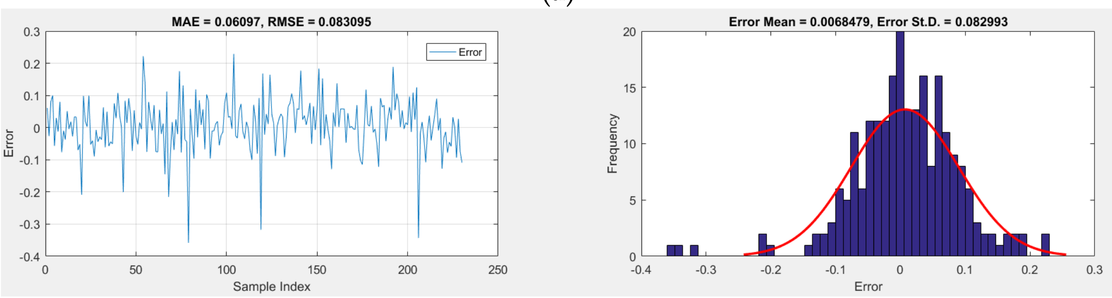


(**b**)

Figure 8. Frequency and minimum value of errors in HHO−ANFIS−300 best fit structure. (**a**) Training dataset. (**b**) Testing dataset.



(**a**)



(**b**)

Figure 9. Frequency and minimum value of errors in HHO−ANFIS−400 best fit structure. (**a**) Training dataset. (**b**) Testing dataset.

As a result of analyzing and evaluating the following figures that represent the initial testing and training phases of the HHO−ANFIS network for five population sizes (100, 200, 300, 400, and 500), one can conclude that the suggested approach has been adequately tested using the target data. If the network is trained accurately, the design is good, and the amount of error varies with the type of data used for initial validation and testing processes. As a result, the final model will be capable of analyzing and predicting new data, as well as unknown data.

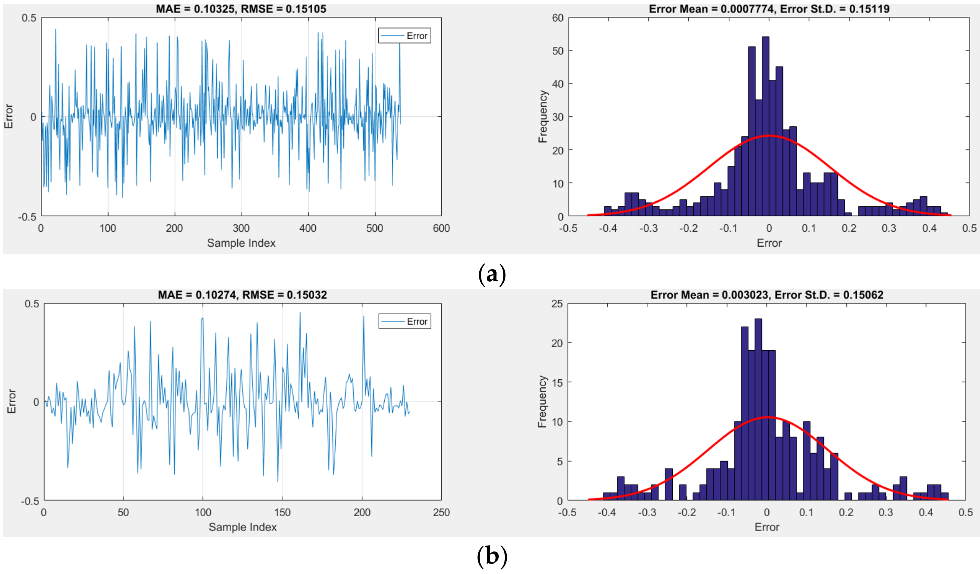


Figure 10. Frequency and minimum value of errors in HHO−ANFIS−500 best fit structure. (**a**) Training dataset. (**b**) Testing dataset.

4.4. Discussion

The majority of reviews and research works focus on the development of machine learning algorithms. Among the research papers studied, there is often a sporadic reporting problem on how machine learning models are developed, with key information missing. The data used in many papers is simply listed without any details about how it was separated for training and testing, for example.

A major limitation in machine learning algorithms lies in the lack of high-quality and real-world datasets and testbeds for evaluating their performance. Even though some datasets are available, such as "UCI machine learning repository" [49], which can be used for training and testing algorithms, high-quality testbeds for evaluating algorithm performance remain scarce. Furthermore, the optimization algorithm performance will become more compelling once tested in a standardized experimental environment.

For future studies on applying machine learning to building load prediction, we expect that the trend will move from the simple application of certain algorithms to a more integrated process and/or framework that incorporates machine learning algorithms training and application as well as data engineering and feature engineering. In particular, it is predicted that the following trends will emerge:

- A shift will be made from algorithm development to data development.
- There will be more and more advances in using domain knowledge to integrate machine learning processes. It is important to note that machine learning development is not just about algorithm development but also the exploration phase of algorithm design for various application scenarios.
- Various machine learning techniques will be implemented in the development of the models, including clustering, active learning, data augmentation, and so on.
- The future of machine learning will feature plug-and-play algorithms and models that require minimal tuning by users.
- As we see it, the major problems related to the future research and investigation of machine learning-based approaches for load prediction are models related to a real building, reduced engineering expenses, and the realization of automation. More specifically:
- The important benefits of machine learning methods for load prediction are automation and engineering cost reduction, but incorporating the algorithms and models into existing BAS and IoT systems can be challenging.

- It is essential to improve algorithms to be applied to different types of buildings, energy systems, and weather to realize automation. A framework based on specific data and buildings can be difficult to automate and tune.
- A lack of real-building testbeds with high data quality also poses a challenge to testing the extendibility of algorithms.

5. Conclusions

The main purpose of this article is to explore the feasibility of estimating building energy performance using a fuzzy approach. It is essential to take into account the building characteristics when determining the amount of heating load required. Therefore, to design and construct energy-efficient buildings and have appropriate indoor conditions, choosing the most related building characteristics, such as those related to heating is helpful. The current study builds on previous research which designed some buildings with the purpose of predicting the heating and cooling systems of buildings. This was done based on the following variables: relative compactness, orientation, roof area, wall area, overall height, surface area, glazing area, and glazing area distribution.

A fuzzy methodology has been studied in this research, i.e., HHO−ANFIS. The training and testing processes were repeated 1000 times in ten population sizes (50, 100, 150, 200, 250, 300, 350, 400, 450, and 500) to the whole dataset, which was randomly permuted in each run prior to splitting it into training and testing subsets. The analysis results show that the HHO−ANFIS with a population size of 400, has a high value of R2 (0.98709 and 0.98794), and a low value of RMSE (0.08769 and 0.08281) in the training and testing dataset, respectively. According to the high value of R2 (98%) and low value of RMSE, HHO−ANFIS can be used in predicting the heating load of residential buildings.

Author Contributions: H.M.: methodology, software; B.L.V.: data curation, revision preparation. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data can be made available based on the reader's request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dadon, I. Planning the Second Generation of Smart Cities: Technology to handle the pressures of urbanization. *IEEE Electrif. Mag.* **2019**, *7*, 6–15. [CrossRef]
2. Pérez-Lombard, L.; Ortiz, J.; Pout, C. A review on buildings energy consumption information. *Energy Build.* **2008**, *40*, 394–398. [CrossRef]
3. Han, Y.; Xu, X.; Zhao, Y.; Wang, X.; Chen, Z.; Liu, J. Impact of consumer preference on the decision-making of prefabricated building developers. *J. Civ. Eng. Manag.* **2022**, *28*, 166–176. [CrossRef]
4. Moradzadeh, A.; Sadeghian, O.; Pourhossein, K.; Mohammadi-Ivatloo, B.; Anvari-Moghaddam, A. Improving Residential Load Disaggregation for Sustainable Development of Energy via Principal Component Analysis. *Sustainability* **2020**, *12*, 3158. [CrossRef]
5. Gao, W.; Alsarraf, J.; Moayedi, H.; Shahsavar, A.; Nguyen, H. Comprehensive preference learning and feature validity for designing energy-efficient residential buildings using machine learning paradigms. *Appl. Soft Comput.* **2019**, *84*, 105748. [CrossRef]
6. Wei, Y.; Zhang, X.; Shi, Y.; Xia, L.; Pan, S.; Wu, J.; Han, M.; Zhao, X. A review of data-driven approaches for prediction and classification of building energy consumption. *Renew. Sustain. Energy Rev.* **2018**, *82*, 1027–1047. [CrossRef]
7. Abd Alla, S.; Bianco, V.; Tagliafico, L.A.; Scarpa, F. Life-cycle approach to the estimation of energy efficiency measures in the buildings sector. *Appl. Energy* **2020**, *264*, 114745. [CrossRef]
8. Zhang, C.; Ali, A.; Sun, L. Investigation on low-cost friction-based isolation systems for masonry building structures: Experimental and numerical studies. *Eng. Struct.* **2021**, *243*, 112645. [CrossRef]
9. Zhao, J.; Liu, X. A hybrid method of dynamic cooling and heating load forecasting for office buildings based on artificial intelligence and regression analysis. *Energy Build.* **2018**, *174*, 293–308. [CrossRef]
10. Di Foggia, G. Energy efficiency measures in buildings for achieving sustainable development goals. *Heliyon* **2018**, *4*, e00953. [CrossRef]

11. Shi, E.; Jabari, F.; Anvari-Moghaddam, A.; Mohammadpourfard, M.; Mohammadi-Ivatloo, B. Risk-Constrained Optimal Chiller Loading Strategy Using Information Gap Decision Theory. *Appl. Sci.* **2019**, *9*, 1925. [CrossRef]
12. Nebot, À.; Mugica, F. Energy Performance Forecasting of Residential Buildings Using Fuzzy Approaches. *Appl. Sci.* **2020**, *10*, 720. [CrossRef]
13. Moayedi, H.; Mosavi, A. Double-Target Based Neural Networks in Predicting Energy Consumption in Residential Buildings. *Energies* **2021**, *14*, 1331. [CrossRef]
14. Yang, L.; Yan, H.; Lam, J.C. Thermal comfort and building energy consumption implications—A review. *Appl. Energy* **2014**, *115*, 164–173. [CrossRef]
15. Mansour-Saatloo, A.; Agabalaye-Rahvar, M.; Mirzaei, M.A.; Mohammadi-Ivatloo, B.; Abapour, M.; Zare, K. Robust scheduling of hydrogen based smart micro energy hub with integrated demand response. *J. Clean. Prod.* **2020**, *267*, 122041. [CrossRef]
16. Mirzaei, M.A.; Yazdankhah, A.S.; Mohammadi-Ivatloo, B. Stochastic security-constrained operation of wind and hydrogen energy storage systems integrated with price-based demand response. *Int. J. Hydrogen Energy* **2018**, *44*, 14217–14227. [CrossRef]
17. Jozi, A.; Pinto, T.; Praça, I.; Vale, Z. Decision Support Application for Energy Consumption Forecasting. *Appl. Sci.* **2019**, *9*, 699. [CrossRef]
18. Zhou, Y.; Wu, J.; Long, C. Evaluation of peer-to-peer energy sharing mechanisms based on a multiagent simulation framework. *Appl. Energy* **2018**, *222*, 993–1022. [CrossRef]
19. Cui, S.; Wang, Y.-W.; Xiao, J.-W. Peer-to-Peer Energy Sharing Among Smart Energy Buildings by Distributed Transaction. *IEEE Trans. Smart Grid* **2019**, *10*, 6491–6501. [CrossRef]
20. Kneifel, J. Life-cycle carbon and cost analysis of energy efficiency measures in new commercial buildings. *Energy Build.* **2010**, *42*, 333–340. [CrossRef]
21. Penna, P.; Prada, A.; Cappelletti, F.; Gasparella, A. Multi-objectives optimization of Energy Efficiency Measures in existing buildings. *Energy Build.* **2015**, *95*, 57–69. [CrossRef]
22. Macas, M.; Moretti, F.; Fonti, A.; Giantomassi, A.; Comodi, G.; Annunziato, M.; Pizzuti, S.; Capra, A. The role of data sample size and dimensionality in neural network based forecasting of building heating related variables. *Energy Build.* **2016**, *111*, 299–310. [CrossRef]
23. Zhao, Y.; Foong, L.K. Predicting electrical power output of combined cycle power plants using a novel artificial neural network optimized by electrostatic discharge algorithm. *Measurement* **2022**, *198*, 111405. [CrossRef]
24. Foong, L.K.; Zhao, Y.; Bai, C.; Xu, C. Efficient metaheuristic-retrofitted techniques for concrete slump simulation. Smart Structures and Systems. *Int. J.* **2021**, *27*, 745–759.
25. Zhao, Y.; Yan, Q.; Yang, Z.; Yu, X.; Jia, B. A Novel Artificial Bee Colony Algorithm for Structural Damage Detection. *Adv. Civ. Eng.* **2020**, *2020*, 3743089. [CrossRef]
26. Zhao, Y.; Hu, H.; Bai, L.; Tang, M.; Chen, H.; Su, D. Fragility Analyses of Bridge Structures Using the Logarithmic Piecewise Function-Based Probabilistic Seismic Demand Model. *Sustainability* **2021**, *13*, 7814. [CrossRef]
27. Jing, Z.; Cai, M.; Pipattanasomporn, M.; Rahman, S.; Kothandaraman, R.; Malekpour, A.; Paaso, E.A.; Bahramirad, S. Commercial building load forecasts with artificial neural network. In Proceedings of the 2019 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), Washington, DC, USA, 18–21 February 2019.
28. Zhang, L.; Wen, J. A systematic feature selection procedure for short-term data-driven building energy forecasting model development. *Energy Build.* **2018**, *183*, 428–442. [CrossRef]
29. Le, L.T.; Nguyen, H.; Dou, J.; Zhou, J. A comparative study of PSO-ANN, GA-ANN, ICA-ANN, and ABC-ANN in estimating the heating load of buildings' energy efficiency for smart city planning. *Appl. Sci.* **2019**, *9*, 2630. [CrossRef]
30. Kwok, S.S.; Lee, E.W. A study of the importance of occupancy to building cooling load in prediction by intelligent approach. *Energy Convers. Manag.* **2011**, *52*, 2555–2564. [CrossRef]
31. Chaudhuri, T.; Soh, Y.C.; Li, H.; Xie, L. A feedforward neural network based indoor-climate control framework for thermal comfort and energy saving in buildings. *Appl. Energy* **2019**, *248*, 44–53. [CrossRef]
32. Yu, Z.; Haghighat, F.; Fung, B.C.M.; Yoshino, H. A decision tree method for building energy demand modeling. *Energy Build.* **2010**, *42*, 1637–1646. [CrossRef]
33. Roy, S.S.; Samui, P.; Nagtode, I.; Jain, H.; Shivaramakrishnan, V.; Mohammadi-Ivatloo, B. Forecasting heating and cooling loads of buildings: A comparative performance analysis. *J. Ambient. Intell. Humaniz. Comput.* **2019**, *11*, 1253–1264. [CrossRef]
34. Zhao, Y.; Wang, Z. Subset simulation with adaptable intermediate failure probability for robust reliability analysis: An unsupervised learning-based approach. *Struct. Multidiscip. Optim.* **2022**, *65*, 172. [CrossRef]
35. Chou, J.-S.; Bui, D.-K. Modeling heating and cooling loads by artificial intelligence for energy-efficient building design. *Energy Build.* **2014**, *82*, 437–446. [CrossRef]
36. Bacher, P.; Madsen, H.; Nielsen, H.A.; Perers, B. Short-term heat load forecasting for single family houses. *Energy Build.* **2013**, *65*, 101–112. [CrossRef]
37. Yu, J.; Bae, M.; Bang, H.-C.; Kim, S.-J. Cloud-based building management systems using short-term cooling load forecasting. In Proceedings of the 2013 IEEE Globecom Workshops (GC Wkshps), Atlanta, GA, USA, 9–13 December 2013; pp. 896–900.
38. Favre, L.; Schafer, T.M.; Robyr, J.-L.; Niederhauser, E.-L. Intelligent algorithm for energy, both thermal and electrical, economic and ecological optimization for a smart building. In Proceedings of the 2018 IEEE International Energy Conference (ENERGYCON), Limassol, Cyprus, 3–7 June 2018. [CrossRef]

39. Ben-Nakhi, A.E.; Mahmoud, M.A. Cooling load prediction for buildings using general regression neural networks. *Energy Convers. Manag.* **2004**, *45*, 2127–2141. [CrossRef]
40. Dounis, A.; Caraiscos, C. Advanced control systems engineering for energy and comfort management in a building environment—A review. *Renew. Sustain. Energy Rev.* **2009**, *13*, 1246–1261. [CrossRef]
41. Han, Y.; Yan, X.; Piroozfar, P. An overall review of research on prefabricated construction supply chain management. *Eng. Constr. Arch. Manag.* 2022, *ahead-of-print*. [CrossRef]
42. Lee, W.-S. Evaluating and ranking energy performance of office buildings using fuzzy measure and fuzzy integral. *Energy Convers. Manag.* **2010**, *51*, 197–203. [CrossRef]
43. Zheng, W.; Liu, X.; Yin, L. Sentence Representation Method Based on Multi-Layer Semantic Network. *Appl. Sci.* **2021**, *11*, 1316. [CrossRef]
44. Zhao, Y.; Joseph, A.J.J.M.; Zhang, Z.; Ma, C.; Gul, D.; Schellenberg, A.; Hu, N. Deterministic snap-through buckling and energy trapping in axially-loaded notched strips for compliant building blocks. *Smart Mater. Struct.* **2019**, *29*, 02LT03. [CrossRef]
45. Chen, S.; Zhang, H.; Wang, L.; Yuan, C.; Meng, X.; Yang, G.; Shen, Y.; Lu, Y. Experimental study on the impact disturbance damage of weakly cemented rock based on fractal characteristics and energy dissipation regulation. *Theor. Appl. Fract. Mech.* **2022**, *122*, 103665. [CrossRef]
46. Bui, X.-N.; Moayedi, H.; Rashid, A.S.A. Developing a predictive method based on optimized M5Rules–GA predicting heating load of an energy-efficient building system. *Eng. Comput.* **2019**, *36*, 931–940. [CrossRef]
47. Nebot, À.; Mugica, F. Fuzzy Approaches Improve Predictions of Energy Performance of Buildings. *Simultech* **2013**, *1*, 504–511. [CrossRef]
48. Tsanas, A.; Xifara, A. Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. *Energy Build.* **2012**, *49*, 560–567. [CrossRef]
49. Set, I.D. Pima+ Indians+ Diabetes. 2007. Available online: http://archive.ics.uci.edu/ml/datasets (accessed on 10 September 2014).
50. Ikram, R.M.A.; Dai, H.-L.; Ewees, A.A.; Shiri, J.; Kisi, O.; Zounemat-Kermani, M. Application of improved version of multi verse optimizer algorithm for modeling solar radiation. *Energy Rep.* **2022**, *8*, 12063–12080. [CrossRef]
51. Cheng, M.-Y.; Cao, M.-T. Accurately predicting building energy performance using evolutionary multivariate adaptive regression splines. *Appl. Soft Comput.* **2014**, *22*, 178–188. [CrossRef]
52. Castelli, M.; Trujillo, L.; Vanneschi, L.; Popovič, A. Prediction of energy performance of residential buildings: A genetic programming approach. *Energy Build.* **2015**, *102*, 67–74. [CrossRef]
53. Ikram, R.M.A.; Ewees, A.A.; Parmar, K.S.; Yaseen, Z.M.; Shahid, S.; Kisi, O. The viability of extended marine predators algorithm-based artificial neural networks for streamflow prediction. *Appl. Soft Comput.* **2022**, *131*, 109739. [CrossRef]
54. Zheng, W.; Zhou, Y.; Liu, S.; Tian, J.; Yang, B.; Yin, L. A Deep Fusion Matching Network Semantic Reasoning Model. *Appl. Sci.* **2022**, *12*, 3416. [CrossRef]
55. Moayedi, H.; Bui, D.T.; Dounis, A.; Lyu, Z.; Foong, L.K. Predicting Heating Load in Energy-Efficient Buildings Through Machine Learning Techniques. *Appl. Sci.* **2019**, *9*, 4338. [CrossRef]
56. Moayedi, H.; Nazir, R.; Gör, M.; Kassim, K.A.; Foong, L.K. A new real-time monitoring technique in calculation of the p-y curve of single thin steel piles considering the influence of driven energy and using strain gauge sensors. *Measurement* **2019**, *153*, 107365. [CrossRef]
57. Chen, J.; Tong, H.; Yuan, J.; Fang, Y.; Gu, R. Permeability Prediction Model Modified on Kozeny-Carman for Building Foundation of Clay Soil. *Buildings* **2022**, *12*, 1798. [CrossRef]
58. Nauck, D.; Kruse, R. Neuro-fuzzy systems for function approximation. *Fuzzy Sets Syst.* **1999**, *101*, 261–271. [CrossRef]
59. Zhao, Y.; Moayedi, H.; Bahiraei, M.; Foong, L.K. Employing TLBO and SCE for optimal prediction of the compressive strength of concrete. *Smart Struct. Syst.* **2020**, *26*, 753–763.
60. Gu, M.; Cai, X.; Fu, Q.; Li, H.; Wang, X.; Mao, B. Numerical Analysis of Passive Piles under Surcharge Load in Extensively Deep Soft Soil. *Buildings* **2022**, *12*, 1988. [CrossRef]
61. Zheng, W.; Tian, X.; Yang, B.; Liu, S.; Ding, Y.; Tian, J.; Yin, L. A Few Shot Classification Methods Based on Multiscale Relational Networks. *Appl. Sci.* **2022**, *12*, 4059. [CrossRef]
62. Zheng, W.; Liu, X.; Ni, X.; Yin, L.; Yang, B. Improving Visual Reasoning Through Semantic Representation. *IEEE Access* **2021**, *9*, 91476–91486. [CrossRef]
63. Moayedi, H.; Varamini, N.; Mosallanezhad, M.; Foong, L.K.; Le, B.N. Applicability and comparison of four nature-inspired hybrid techniques in predicting driven piles' friction capacity. *Transp. Geotech.* **2022**, *37*, 100875. [CrossRef]
64. Yan, B.; Ma, C.; Zhao, Y.; Hu, N.; Guo, L. Geometrically Enabled Soft Electroactuators via Laser Cutting. *Adv. Eng. Mater.* **2019**, *21*, 1900664. [CrossRef]
65. Zheng, W.; Yin, L.; Chen, X.; Ma, Z.; Liu, S.; Yang, B. Knowledge base graph embedding module design for Visual question answering model. *Pattern Recognit.* **2021**, *120*, 108153. [CrossRef]
66. Wu, P.; Liu, A.; Fu, J.; Ye, X.; Zhao, Y. Autonomous surface crack identification of concrete structures based on an improved one-stage object detection algorithm. *Eng. Struct.* **2022**, *272*, 114962. [CrossRef]
67. Heidari, A.A.; Mirjalili, S.; Faris, H.; Aljarah, I.; Mafarja, M.; Chen, H. Harris hawks optimization: Algorithm and applications. *Futur. Gener. Comput. Syst.* **2019**, *97*, 849–872. [CrossRef]

68. Zhao, Y.; Zhong, X.; Foong, L.K. Predicting the splitting tensile strength of concrete using an equilibrium optimization model. Steel and Composite Structures. *Int. J.* **2021**, *39*, 81–93.
69. Yang, X.-S. *Nature-Inspired Metaheuristic Algorithms*; Luniver Press: Beckington, UK, 2010.